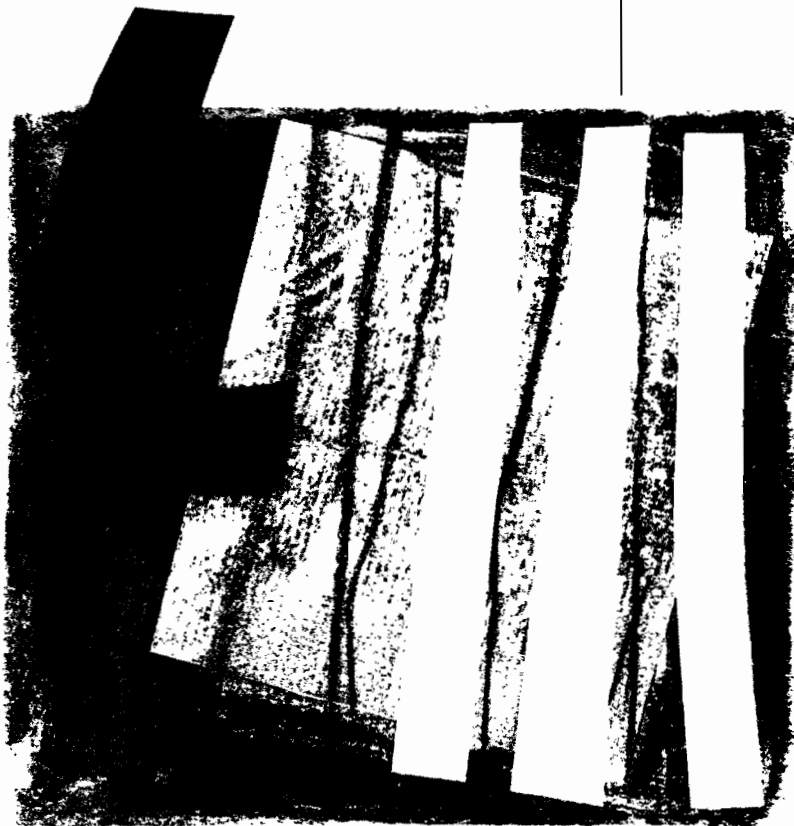


**A POWERFUL PORTMANTEAU TEST
OF LACK OF FIT FOR TIME SERIES**

**Daniel Peña
Julio Rodríguez**

00-82



WORKING PAPERS

A POWERFUL PORTMANTEAU TEST OF LACK OF FIT FOR TIME SERIES

Daniel Peña y Julio Rodríguez*

Abstract

A new portmanteau test for time series more powerful than the tests of Ljung and Box (1978) and Monti (1994) is proposed. The test is based on the p th root of the determinant of the p th autocorrelation matrix. It is shown that this statistic can be interpreted as the geometric mean of the squared multiple correlation coefficients with m lag values when m goes from 1 to p . It can also be interpreted as a geometric mean of the partial autocorrelation coefficients. The asymptotic distribution of the test statistic is obtained. This distribution is a linear combination of chi-squared distributions and it is shown that it can be approximated by a gamma distribution. The power of the test is compared with that of the Ljung and Box and Monti tests and it is shown that the proposed test can be up to 50% more powerful depending upon the model and sample size.

The test is applied to the detection of nonlinearity by using the same matrix but with coefficients that are now the autocorrelations of the squared residuals. The new test is more powerful than the test of McLeod and Li (1983) for nonlinearity. An example is presented in which this test detects nonlinearity in the residuals of the sunspot series.

Keywords: Autocorrelation; dependency coefficient; nonlinearity test.

*Peña, Department of Statistics and Econometrics, Universidad Carlos III de Madrid, e-mail: dpena@est-econ.uc3m.es; Rodríguez, Department of Statistics and Econometrics, Universidad Carlos III de Madrid, e-mail: puerta@est-econ.uc3m.es

A Powerful Portmanteau Test of Lack of Fit for Time Series

Daniel Peña and Julio Rodríguez*

November 30, 2000

Abstract

A new portmanteau test for time series more powerful than the tests of Ljung and Box (1978) and Monti (1994) is proposed. The test is based on the p th root of the determinant of the p th autocorrelation matrix. It is shown that this statistic can be interpreted as the geometric mean of the squared multiple correlation coefficients with m lag values when m goes from 1 to p . It can also be interpreted as a geometric mean of the partial autocorrelation coefficients. The asymptotic distribution of the test statistic is obtained. This distribution is a linear combination of chi-squared distributions and it is shown that it can be approximated by a gamma distribution. The power of the test is compared with that of the Ljung and Box and Monti tests and it is shown that the proposed test can be up to 50% more powerful depending upon the model and sample size.

The test is applied to the detection of nonlinearity by using the same matrix but with coefficients that are now the autocorrelations of the squared residuals. The new test is more powerful than the test of McLeod and Li (1983) for nonlinearity. An example is presented in which this test detects nonlinearity in the residuals of the sunspot series.

KEY WORDS: Autocorrelation; Dependency coefficient; Nonlinearity test.

1 INTRODUCTION

Suppose a time series $\{X_t\}$ generated by a stationary and invertible ARMA(p, q) process of the form $\phi(B)X_t = \theta(B)\varepsilon_t$, where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and $\phi(B)$ and $\theta(B)$ are polynomials given by $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$, where $B^k X_t = X_{t-k}$. Usually X_t is some transformation of an observed time series such as differencing. The residuals of this model are defined by $\hat{\varepsilon}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)X_t$ where $\hat{\theta}$, $\hat{\phi}$ and $\hat{\sigma}_\varepsilon^2$ are the maximum likelihood estimators of $\phi = (\phi_1, \dots, \phi_p)'$, $\theta = (\theta_1, \dots, \theta_q)'$ and σ_ε^2 , respectively. Several diagnostic goodness of fit tests for ARIMA models have been proposed based on the residual autocorrelation coefficients given by

$$\hat{r}_k = \frac{\sum_{t=k+1}^n \hat{\varepsilon}_t \hat{\varepsilon}_{t-k}}{\sum_{t=1}^n \hat{\varepsilon}_t^2} \quad (k = 1, 2, \dots).$$

Box and Pierce (1970) proposed a portmanteau test to check the adequacy of the fitted model using the statistic $Q = n \sum_{k=1}^m \hat{r}_k^2$, and they showed that the asymptotic distribution of Q can be approximated by a χ^2 distribution with $m - (p + q)$ degrees of freedom. In order to improve the

*Department of Statistics and Econometrics, University Carlos III of Madrid, Calle Madrid 126, CP: 28903, Getafe, Madrid, Spain, (email: dpena@est-econ.uc3m.es - puerta@est-econ.uc3m.es)

χ^2 approximation, Ljung and Box (1978) proposed to replace the autocorrelation coefficients $hat{r}_k$ by the standardized coefficients, $\tilde{r}_i = [(n+2)/(n-i)]^{1/2} \hat{r}_i$, leading to the statistic

$$Q_{LB} = n(n+2) \sum_{k=1}^m (n-k)^{-1} \hat{r}_k^2. \quad (1)$$

When m is small, Ljung (1986) showed that the estimated size of Q_{LB} can be improved using the scaled χ^2 distribution and she noted that using too many residual autocorrelations can reduce the power of the test. More recently, Monti (1994) proposed a portmanteau test of goodness of fit based on the residual partial autocorrelations. Let $\hat{\pi}_k$ be the k th residual partial autocorrelation, i.e. the correlation between the regression residuals $\hat{\varepsilon}_t$ and $\hat{\varepsilon}_{t-k}$, holding $\hat{\varepsilon}_{t-1}, \dots, \hat{\varepsilon}_{t-k+1}$ 'fixed'. If the model is correctly specified, $\hat{\pi}_k$ is approximately distributed as a normal random variable with mean zero and variance $(n-k)/(n(n+2))$. Thus, a portmanteau test statistic similar to the Ljung-Box statistic can be written using the partial autocorrelation coefficients:

$$Q_{MT} = n(n+2) \sum_{k=1}^m (n-k)^{-1} \hat{\pi}_k^2. \quad (2)$$

Under the assumption that the time series has been generated by an ARMA(p, q) the asymptotic distribution of Q_{MT} is χ^2 with $m - (p+q)$ degrees of freedom. Monti (1994) showed by simulations that the estimated size of the test is adequate with moderate sample sizes and the test is more powerful than the Ljung-Box test when the fitted model underestimates the order of the moving average component. Kwan and Wu (1997) examined the finite-sample properties of both tests, (1) and (2), via Monte Carlo simulation. They computed the empirical powers of Q_{MT} and Q_{LB} when the data were generated with monthly seasonality, finding only slight differences between Q_{MT} and Q_{LB} .

This paper proposes a new portmanteau goodness of fit test based on a measure of multivariate dependency and is organized as follows. Section 2 presents the test and its main properties. Section 3 presents the asymptotic distribution of the test statistic. Section 4 includes a Monte Carlo study of the properties of the test. Section 5 extends the procedure and propose a test for nonlinearity using the autocorrelations of the squared residuals.

2 THE PROPOSED TEST

The estimated residuals can be considered as a sample of multivariate data from some distribution and we are interested in testing if the covariance matrix of their distribution is proportional to the identity, or equivalently, if the correlation matrix is the identity. In multivariate analysis the likelihood ratio test for checking if a set of random variables has a scalar covariance matrix is proportional to the determinant of the correlation matrix of the multivariate variables, $|R|$. Thus, it is sensible to explore a test based on this statistic.

For stationary time series data the residual correlation matrix of order m , \hat{R}_m is given by

$$\hat{R}_m = \begin{bmatrix} 1 & \hat{r}_1 & \cdots & \hat{r}_m \\ \hat{r}_1 & 1 & \cdots & \hat{r}_{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{r}_m & \hat{r}_{m-1} & \cdots & 1 \end{bmatrix} \quad (3)$$

and we have that

$$|\hat{R}_m| = |\hat{R}_{m-1}| (1 - \hat{R}_i^2)$$

where $\hat{R}_i^2 = \hat{r}'_{(m)} \hat{R}_m^{-1} \hat{r}_{(m)}$, with $\hat{r}_{(m)} = (\hat{r}_1, \dots, \hat{r}_m)'$, is the determination coefficient in the linear fit $\hat{\varepsilon}_t = \sum_{j=1}^i b_j \hat{\varepsilon}_{t-j} + u_t$. By recursive use of this expression we can write

$$|\hat{R}_p| = \prod_{i=1}^m (1 - \hat{R}_i^2) \quad (4)$$

We can define the residuals mean autodependency, $D(\hat{\varepsilon}_t, m)$, as the average value (in the log scale) of the squared correlations obtained when fitting autoregressive models of increasing order to the time series by

$$\log(1 - D(\hat{\varepsilon}_t, m)) = \frac{1}{m} \sum_{i=1}^m \log(1 - \hat{R}_i^2). \quad (5)$$

This statistic has an alternative interpretation as a weighted average of the partial autocorrelation coefficients, π_i . It is well known that

$$(1 - \pi_i^2) = \frac{1 - R_i^2}{1 - R_{i-1}^2},$$

and using this expression in (5) it can be obtained that

$$\log(1 - D(\hat{\varepsilon}_t, m)) = \sum_{i=1}^m w_i \log(1 - \hat{\pi}_i^2), \quad (6)$$

where $w_i = (m + 1 - i)/m$. We propose to test for autocorrelation in the estimated residuals with the portmanteau statistic:

$$D_m(\hat{\varepsilon}_t) = nD(\hat{\varepsilon}_t, m) = n \left[1 - |\hat{R}_m|^{1/m} \right]. \quad (7)$$

3 DISTRIBUTION OF THE PROPOSED STATISTIC

In this section we first obtain the asymptotic distribution of the proposed statistic (7) for all m , where m is the number of sample autocorrelations. As this distribution is complicated, we follow Box and Pierce (1970) in obtaining an approximation of this distribution when m is moderately high. Then, we will show by Monte Carlo simulation that this approximation work well in small samples.

3.1 Asymptotic distribution of $D_m(\hat{\varepsilon}_t)$

The autodependency is a continuous function of the sample partial autocorrelations, $\hat{\pi}_i$, as shown in equation (6). Defining $\hat{\pi}_{(m)} = (\hat{\pi}_1, \dots, \hat{\pi}_m)'$, and using the result by Monti (1994) we have that $n^{1/2} \hat{\pi}_{(m)}$ is asymptotically multivariate normal with zero mean vector and covariance matrix $(I_m - Q_m)$, where $Q_m = X_m V^{-1} X_m'$, V is the information matrix for the parameters ϕ and θ , and X_m is an $m \times (p + q)$ matrix, with elements ϕ' and θ' defined by $1/\phi(B) = \sum_{i=0}^{\infty} \phi'_i B^i$ and $1/\theta(B) = \sum_{i=0}^{\infty} \theta'_i B^i$, (see Brockwell and Davis, 1991, pp. 296-304). The coefficients ϕ'_i and θ'_i are readily computed using the recursive procedure of Box and Jenkins (1976, pp. 132-134).

Theorem 1 *If the model is correctly identified, $D_m(\hat{\varepsilon}_t)$ is asymptotically distributed as $\sum_{i=1}^m \lambda_i \chi_{1,i}^2$, where $\chi_{1,i}^2$ ($i = 1, \dots, m$) are independent χ_1^2 random variables and λ_i ($i = 1, \dots, m$) are the eigenvalues of $(I_m - Q_m)W_m$, where W_m is a diagonal matrix with elements $w_{ii} = (m - i + 1)/m$ ($i = 1, \dots, m$).*

Proof. Suppose that under the null hypothesis, $D_m(\hat{\varepsilon}_t)$ is asymptotically distributed as the random variable X . Then, applying the δ -method (Arnold, 1990) to $g(x) = \log(1 - x)$, it follows that $-n \log |\hat{R}_m|^{1/m}$ is also asymptotically distributed as X . From (6) we obtain the equivalent expression

$$-n \log \left(|\hat{R}_m|^{1/m} \right) = -n \sum_{i=1}^m \frac{m - i + 1}{m} \ln (1 - \hat{\pi}_i^2). \quad (8)$$

To find the distribution of (8), suppose that $(n\hat{\pi}_1^2, n\hat{\pi}_2^2, \dots, n\hat{\pi}_m^2)$ is asymptotically distributed as the random variable Y . Then, applying the multivariate δ -method (Arnold, 1990) to $g(\hat{\pi}_1^2, \dots, \hat{\pi}_m^2) = -\sum_{i=1}^m \frac{m-i+1}{m} \ln (1 - \hat{\pi}_i^2)$, it follows that

$$-n \sum_{i=1}^m \frac{m - i + 1}{m} \ln (1 - \hat{\pi}_i^2) \rightarrow (1, \frac{m-1}{m}, \dots, \frac{1}{m})Y. \quad (9)$$

where \rightarrow stands for convergence in distribution and from the Cramer-Wold theorem (Arnold 1990), it follows that

$$(1, \frac{m-1}{m}, \dots, \frac{1}{m})(n\hat{\pi}_1^2, n\hat{\pi}_2^2, \dots, n\hat{\pi}_m^2)' \rightarrow (1, \frac{m-1}{m}, \dots, \frac{1}{m})Y. \quad (10)$$

Using the fact that $n^{1/2}\hat{\pi}_{(m)}$ is asymptotically distributed as $N(0, I_m - Q_m)$ and from the theorem on quadratic forms given by Box (1954), it follows that

$$(1, \frac{m-1}{m}, \dots, \frac{1}{m})(n\hat{\pi}_1^2, n\hat{\pi}_2^2, \dots, n\hat{\pi}_m^2)' = n\hat{\pi}_i' D_m \hat{\pi}_i \rightarrow \sum_{i=1}^m \lambda_i \chi_{1,i}^2. \quad (11)$$

Finally, from (10) and (11)

$$(1, \frac{m-1}{m}, \dots, \frac{1}{m})Y \rightarrow \sum_{i=1}^m \lambda_i \chi_{1,i}^2.$$

and from (9)

$$D_m(\hat{\varepsilon}) \rightarrow \sum_{i=1}^m \lambda_i \chi_{1,i}^2. \quad (12)$$

■

For a general ARMA model the expression for the eigenvalues of $(I_m - Q_m)W_m$ is complicated. However, the eigenvalues are readily calculated for any given ϕ , θ and m . In the following subsection, we propose an approximation of the distribution (12) for a moderately high m .

3.2 Approximation of the distribution $D_m(\hat{\varepsilon}_t)$

The probability $pr(D_m(\hat{\varepsilon}) > x)$ can be evaluated by inverting the characteristic function of $\sum_{i=1}^m \lambda_i \chi_{1,i}^2$ (Imhof 1961). The procedure requires only one-dimensional numerical integration, but for simplicity, we prefer to use the approximation proposed by Box (1954). He suggested approaching a distribution of the form (12) by a distribution of the form $a\chi_b^2$, with mean and variance equal to

those of the exact distribution. Following this suggestion, we approximate the distribution (12) with parameters a and b defined by $a = \sum \lambda_i^2 / \sum \lambda_i$ and $b = (\sum \lambda_i)^2 / \sum \lambda_i^2$, where λ_i are the eigenvalues of the matrix $(I_m - Q_m)W_m$.

Box and Pierce (1970) and McLeod (1978) approximate the matrix $Q_m = X_m V^{-1} X_m'$ by the projection matrix $Q_m = X_m (X_m' X_m)^{-1} X_m'$ when m is moderately high. This approximation is useful for computing an expression for a and b that does not depend on the ARMA parameters ϕ and θ . Then we have

$$\sum_{i=1}^m \lambda_i = \text{tr}((I_m - Q_m) W_m) = \text{tr}(W_m) - \text{tr}(Q_m) + (1/m) \text{tr}(Q_m C_m) \quad (13)$$

where C_m is a diagonal matrix with elements $c_{ii} = i$, $i = 0, \dots, (m-1)$, and

$$\begin{aligned} \sum_{i=1}^m \lambda_i^2 &= \text{tr}((I_m - Q_m) W_m^2) = \text{tr}(W_m^2) - \text{tr}(Q_m) + \\ &+ (2/m) \text{tr}(Q_m C_m) - (1/m^2) \text{tr}(Q_m C_m^2). \end{aligned} \quad (14)$$

An alternative expression for $\sum \lambda_i$ and $\sum \lambda_i^2$ can be obtained using the Cholesky decomposition of the matrix $(I_m - Q_m)$, see Velilla (1994). Using the fact that Q_m is an idempotent matrix with rank $p + q$, then (13) and (14) can be written as a function of p , q , m and q_{ii} , where the q_{ii} are the elements in the diagonal of Q_m ,

$$\sum_{i=1}^m \lambda_i = \frac{m+1}{2} - (p+q) + \frac{1}{m} \sum_{i=2}^m (i-1) q_{ii}, \quad (15)$$

$$\begin{aligned} \sum_{i=1}^m \lambda_i^2 &= \frac{1}{6m} (m+1) (2m+1) - (p+q) + \\ &+ \frac{2}{m} \sum_{i=2}^m (i-1) q_{ii} - \frac{1}{m^2} \sum_{i=2}^m (i-1)^2 q_{ii}. \end{aligned} \quad (16)$$

Now we will show that the terms $(1/m) \sum_{i=2}^m (i-1) q_{ii}$ and $(1/m^2) \sum_{i=2}^m (i-1)^2 q_{ii}$ tend to zero when m increases. As Q_m is an idempotent matrix of rank $p + q$, then $\sum_{i=1}^m q_{ii} = p + q$. Consider the sequences $a_i = i$ and $b_i = (i-1) q_{ii}$. Then $\sum_{i=1}^m (i-1) q_{ii} / i \leq \sum_{i=1}^m q_{ii} = p + q < \infty$ as $m \rightarrow \infty$. Then by Kronecker's lemma (Davidson, 1997, pp. 34-35) we obtain that $(1/m) \sum_{i=1}^m (i-1) q_{ii} \rightarrow 0$. A similar argument is used to show that $(1/m^2) \sum_{i=1}^m (i-1)^2 q_{ii} \rightarrow 0$. Thus, for large m we can approximate equations (15) and (16) by

$$\sum_{i=1}^m \lambda_i = \frac{m+1}{2} - (p+q), \quad (17)$$

$$\sum_{i=1}^m \lambda_i^2 = \frac{1}{6m} (m+1) (2m+1) - (p+q). \quad (18)$$

This approximation of the distribution (12) is stated in the next corollary.

Corollary 2 *If ARMA model is correctly identified, when m is moderately high, the distribution of $D_m(\hat{\epsilon}_t)$ can be approximated by a distribution $a\chi_b^2$ with mean $(m+1)/2 - (p+q)$ and variance $(m+1)(2m+1)/3m - 2(p+q)$.*

3.3 Comparison between the empirical distribution and the approximation of the distribution of $D_m(\hat{\varepsilon}_t)$

In this section we present the results of a Monte Carlo experiment which was carried out to check the approximation to the asymptotic distribution of $D_m(\hat{\varepsilon})$ indicated in Corollary 2. As the χ^2 approximation to the asymptotic distribution improves when using the standardized autocorrelation coefficients estimator $\tilde{r}_{(m)} = (\tilde{r}_1, \dots, \tilde{r}_m)'$, where $\tilde{r}_i = [(n+2)/(n-i)]^{1/2} \hat{r}_i$, we will consider in addition to $D_m(\hat{\varepsilon}_t)$, the statistic based on $\hat{r}_{(m)}$, the statistic $D_m^*(\hat{\varepsilon}_t)$ based on $\tilde{r}_{(m)}$. Under the null hypothesis, $D_m(\hat{\varepsilon}_t)$ and $D_m^*(\hat{\varepsilon}_t)$ are asymptotically equivalent, but obviously, some differences may be expected in their small-sample behaviours. We have included in the study the Q_{LB} and Q_{MT} statistics in order to compare the approximation we are proposing to those used for these statistics. The sample sizes used are 50, 100 and 500. We report here only the results when the generating process is white noise. Similar results have been found when the generating process is a low order AR or MA model and are available upon request from the authors.

[TABLE 1 ABOUT HERE]

[TABLE 2 ABOUT HERE]

In the white noise case, $p = q = 0$, $D_m(a_t)$ can be approximated by $a\chi_b^2$ where $a = (2m+1)/3m$ and $b = 3m(m+1)/(4m+2)$ and it has asymptotic mean and variance $(m+1)/2$ and $(2m+1)(m+1)/3m$. Tables 1 and 2 show the mean and the variance of the empirical distribution of the statistics D_m , D_m^* , Q_{LB} and Q_{MT} , obtained by generating 10,000 replications of Gaussian white noise with sample size 50, 100 and 500 and computing the four statistics in each replication for lags 7, 15 and 20. The last row of Table 1 presents the asymptotic mean and the last row of Table 2 the asymptotic variance. The approximation is quite good in all four cases. In Table 2 we observe that in all cases the variance of D_m^* is closer to the asymptotic variance than that of D_m . The approximation in the distribution of D_m^* works better than in the case of Q_{LB} , where the empirical variance is clearly overestimated.

[FIGURE 1 ABOUT HERE]

Figure 1 illustrates the accuracy of the empirical distribution of D_m and D_m^* to the distribution (12), with 10,000 replications of sample size 100 by an AR(1) process with $\phi_1 = .5$, and $m = 20$. The approximation by the standardized estimators is better than the non-standardized estimators, and based on these simulations, we recommend use of D_m^* specially, for small sample size.

4 COMPARISON OF SIGNIFICANCE LEVELS AND POWER IN SMALL-SAMPLE OF $D_m^*(\hat{\varepsilon}_t)$

[TABLE 3 ABOUT HERE]

In this section we present a comparative study of significance level and power for the statistics D_m^* , Q_{LB} and Q_{MT} . The significance levels of Q_{LB} and Q_{MT} are obtained using the percentiles of the χ^2 distribution, and those of D_m^* , have been obtained by using the approximation of (12) by the $a\chi_b^2$ with a and b obtained from (17) and (18). These significance levels are evaluated under both low order AR and MA models but we report here only the results for the AR(1). Similar results were found in the other cases and they are available upon request from the authors. In each case, 10,000 Gaussian series of 100 observations were generated and the performance of D_m^* was compared to that of the Ljung-Box and Monti statistics. Three values for m , 10, 15 and 20 were considered. Table 3 shows the significance levels, under the AR(1) model, when the nominal levels α are 0.05 and 0.01. The approximation seems to be satisfactory for the three statistics, and for Q_{LB} and Q_{MT} is similar to the results presented by Monti (1994). For $\alpha = 0.05$, the significance

level of D_m^* is between the significance level of Q_{LB} and Q_{MT} . The significance level of D_m^* does not seem to be affected by the value of m considered.

[TABLE 4 ABOUT HERE]

The power is analyzed for the same models as proposed by Monti (1994) to compare the performance of the Q_{MT} and Q_{LB} statistics. Table 4 shows the power of the three tests when, erroneously, an AR(1) or a MA(1) model is fitted to the data. Twenty four different ARMA(2,2) models are considered. In each case 1,000 series of 100 observations were generated and the power was computed with $m = 10$ and $m = 20$. The power decreases as m grows in the three tests, but the loss of power in D_m^* is relatively lower. The power of D_m^* is always the highest and the increase in power with respect to the best of Q_{MT} and Q_{LB} can in some cases be as high as 50% (see models 1, 11 and 23).

5 CHECKING THE LINEARITY ASSUMPTION

Granger and Anderson (1978) suggested that looking at the autocorrelation function of y_t^2 could be useful in identifying nonlinear time series. Maravall (1983) showed that the autocorrelation function, (ACF), of the squared residuals, provides a convenient tool for checking the linearity assumption. If the residuals $\hat{\varepsilon}_t$ are independent so the $\hat{\varepsilon}_t^2$ will be, but if the model is nonlinear and the residuals $\hat{\varepsilon}_t$ are not independent, this feature can appear in the ACF of $\hat{\varepsilon}_t^2$, which is expected to be different from the ACF of a white noise process. McLeod and Li (1983) proposed to detect non linearity in time series data by computing the statistic Q_{LB} on the autocorrelations of the squared residuals from a linear fit. The significance levels are based on the asymptotic χ_m^2 distribution of the test statistic when the process is linear. These same ideas could be applied to the statistic proposed by Monti (1994) and to the new statistic (7) proposed in this article. Then, the proposed nonlinearity test (7) can be written as

$$D_m(\hat{\varepsilon}_t^2) = nD(\hat{\varepsilon}_t^2, m),$$

where the autocorrelation function of $\hat{\varepsilon}_t^2$ is estimated by

$$\hat{r}_k(\hat{\varepsilon}_t^2) = \frac{\sum_{t=k+1}^n (\hat{\varepsilon}_t^2 - \hat{\sigma}^2) (\hat{\varepsilon}_{t-k}^2 - \hat{\sigma}^2)}{\sum_{t=1}^n (\hat{\varepsilon}_t^2 - \hat{\sigma}^2)^2} \quad (k = 1, 2, \dots, m).$$

where $\hat{\sigma}^2 = \sum \hat{\varepsilon}_t^2 / n$, and analogously for Monti's tests $Q_{MT}(\hat{\varepsilon}_t^2)$. In the next lemma we calculate the asymptotic distribution of the statistic $D_m(\hat{\varepsilon}_t^2)$ when ARMA model is correctly identified.

Lemma 3 *If the ARMA model is correctly identified then $D_m(\hat{\varepsilon}_t^2)$ is asymptotically distributed as $\sum_{i=1}^m w_{ii} \chi_{1,i}^2$ where $\chi_{1,i}^2$ ($i = 1, \dots, m$) are independent χ_1^2 random variables and $w_{ii} = (m-i+1)/m$ ($i = 1, \dots, m$).*

Proof. This is based on the result by McLeod and Li (1983) for the distribution of $n^{1/2} \hat{r}_{(m)}(\hat{\varepsilon}_t^2)$, which is $N(0, I_m)$. Applying this result to that obtained by Monti (1994), the asymptotic distribution of $n^{1/2} \hat{\pi}_{(m)}(\hat{\varepsilon}_t^2)$ is $N(0, I_m)$. Following the same reasoning as in Theorem 1, the asymptotic distribution for $Q_D(\hat{\varepsilon}_t^2)$ is obtained. ■

The asymptotic distribution of Lemma 3 can be approximated as before by distribution of form $a\chi_b^2$ proposed by Box (1954). The following Corollary gives the details.

Corollary 4 *If ARMA model is correctly identified, then the asymptotic distribution of $D_m(\hat{\varepsilon}_t^2)$ can be approximated by a distribution of form $a\chi_b^2$, where $a = (2m+1)/3m$ and $b = 3m(m+1)/(4m+2)$.*

5.1 Power comparison

Davis and Petrucci (1986) compared the statistics proposed by McLeod and Li (1983) and by Keenan (1985). They concluded that the statistic proposed by Keenan is more powerful than the one proposed by Li and McLeod but that the power of both statistics is poor except for large sample size, $n > 200$. In this section we compare the power of the statistics D_m^* , Q_{LB} and Q_{MT} for testing the linearity hypothesis, when applied to the squared residuals. As earlier, we use the correction of the statistic D_m^* for small samples based on the standardized autocorrelations, $\tilde{r}_k(\hat{\varepsilon}_t^2) = [(n+2)/(n-i)]^{1/2} \hat{r}_k(\hat{\varepsilon}_t^2)$. We will compare the power for the four nonlinear models that Keenan (1985) used in his study:

$$\text{Model 1, } Y_t = e_t - 0.4e_{t-1} + 0.3e_{t-2} + 0.5e_te_{t-2};$$

$$\text{Model 2, } Y_t = e_t - 0.3e_{t-1} + 0.2e_{t-2} + 0.4e_{t-1}e_{t-2} - 0.25e_{t-2}^2;$$

$$\text{Model 3, } Y_t = 0.4Y_{t-1} - 0.3Y_{t-2} + 0.5Y_{t-1}e_{t-1} + e_t;$$

$$\text{Model 4, } Y_t = 0.4Y_{t-1} - 0.3Y_{t-2} + 0.5Y_{t-1}e_{t-1} + 0.8e_{t-1} + e_t,$$

where the e_t 's are independent $N(0,1)$.

[TABLE 5 ABOUT HERE]

Table 5 summarizes the power results. For each model 1,000 replications of sample size $n = 204$ were generated and fitted with an $AR(p)$ model, where p is selected by the AIC (Akaike 1974) with $p \in \{1, 2, 3, 4\}$. The power of the proposed statistic, $D_m^*(\hat{\varepsilon}_t^2)$, is broadly between 6% (Model 4, $m = 7$) and 40% (Model 2, $m = 24$) higher than the ones based on Q_{LB} and Q_{MT} . None of these tests are powerful in handling Model 1, which contains a concurrent nonlinear term e_te_{t-2} . The difficulty in finding the nonlinearity of Model 1 is observed in the study by Tsay (1986) which compares some statistics for checking the nonlinearity over the same four models. From the comparison of Table 5 and the results obtained by Tsay (1986) we conclude that the proposed test $D_m^*(\hat{\varepsilon}_t^2)$ is better than that proposed by Keenan for all lags, $m = 7, 12$ and 24 and for all the models except Model 2. However, the proposed portmanteau test is slightly worse than that proposed by Tsay.

We are interested in checking the behaviour of the proposed statistics in the detection of non linearity in Threshold Autoregressive models (TAR) which are among the most popular non linear time series models in applied research. With this aim 1,000 replications were generated with sample sizes 100, 200 and 500 from the TAR(2) model with 4 regimes proposed by Tiao and Tsay (1994). This model as an alternative to the AR(2) model to capture the structure of the U.S. real GNP series from the first quarter of 1947 to the first quarter of 1991 with a total of 177 observations. The regimes and their economic interpretation are described in Tiao and Tsay (1994).

[TABLE 6 ABOUT HERE]

Table 6 shows the power of the statistics D_m^* , Q_{LB} and Q_{MT} using the residuals autocorrelations $\hat{\varepsilon}_t$, with $\alpha = 0.05$ to find structure in these samples when fitting an AR(2) model. The values shown in the table are close to the nominal level (0.05) confirming the lack of power of these statistics to detect nonlinearity. Table 7 displays the power of these statistics when the squared residuals $\hat{\varepsilon}_t^2$ of the AR(2) fit is used. The power of the three statistics increases as expected and now the power of D_m^* is clearly higher for all sample sizes. Large samples sizes are needed for these types of statistics to be useful and for $n = 500$ the increase of power of D_m^* with respect to the best of Q_{LB} and Q_{MT} goes from 23% to 35%.

[TABLE 7 ABOUT HERE]

5.2 An example with real data

[TABLE 8 ABOUT HERE]

In this section we apply the $D_m^*(\hat{\varepsilon}_t^2)$ statistic to find nonlinearities in the well known Sunspot time series, Box and Jenkins (1976). This time series has been previously studied by various authors who have applied both linear and nonlinear models. A detailed description of some of the different models which have been proposed can be found in Priestley (1989). We have followed Priestley (1989, p. 882) and fitted an AR(9) model to the sample of 246 observations corresponding to the first 246 observations from 1700. The order of the linear model was determined according to Akaike's criterion. The model is estimated by maximum likelihood and statistics D_m^* , Q_{LB} and Q_{MT} , are computed from the residuals. No structure is found using the lags $m = 12$ and 24 and $\alpha = 0.05$. However, when using the statistic $D_m^*(\hat{\varepsilon}_t^2)$ a clear indication of non linear structure is found. Table 8 presents the values of the three portmanteau statistics applied to the squared residuals. To facilitate the comparison columns 3, 5 and 7 show the ratio between the statistic and the percentile corresponding to $\alpha = 0.05$ for each distribution. The results obtained with the statistic $D_m^*(\hat{\varepsilon}_t^2)$ are more conclusive than those with the statistics Q_{LB} and Q_{MT} , for checking the nonlinearity of the time series. The statistic $D_m^*(\hat{\varepsilon}_t^2)$ clearly suggests a nonlinear structure whereas the results from the Ljung-Box and Monti statistics are not decisive.

6 ACKNOWLEDGMENTS

This research has been sponsored by DGES (Spain) under project PB-96-0111 and the Cátedra BBVA de Calidad. We are very grateful to Andrés Alonso for many stimulating discussions and to Mike Wiper for his help with the final draft.

References

- Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Trans. Automat. Control AC 19*, 203–17.
- Arnold, S. F. (1990). *Mathematical Statistics*. Prentice-Hall International, Inc.
- Box, G. E. P. (1954). Some Theorems on Quadratics Forms Applied in the Study of Analysis of Variance Problems I: Effect of the Inequality of Variance in the One-way Classification. *Annals of Mathematical Statistics 25*, 290–302.
- Box, G. E. P. and G. M. Jenkins (1976). *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco.
- Box, G. E. P. and D. A. Pierce (1970). Distribution of Residuals Autocorrelations in Autoregressive-integrated Moving Average Time Series Models. *Journal of the American Statistical Association 65*, 1509–26.
- Brockwell, P. and R. A. Davis (1991). *Time Series: Theory and Methods* (2nd Edn ed.). New York: Springer Verlag.
- Davidson, J. (1997). *Stochastic Limit Theory*. Oxford University Press.
- Davis, G. E. P. and D. A. Petruccielli (1986). Detecting Non-linearity in Time Series. *The Statistician 35*, 271–280.
- Granger, C. W. J. and A. P. Anderson (1978). *An Introduction to Bilinear Time Series Models*. Gttingen, Vandenhoeck & Rpurecht.

- Imhof, J. P. (1961). Computing the Distribution of Quadratics Forms in Normal Variables. *Biometrika* 48, 419–26.
- Keenan, D. M. (1985). A Tukey Nonadditivity-type Test for Time Series Nonlinearity. *Biometrika* 72, 39–44.
- Kwan, A. C. C. and Y. Wu (1997). Further Results on the Finite-sample Distribution of Monti's Portmanteau Test for the Adequacy of an ARMA(p,q) Model. *Biometrika* 84, 733–36.
- Ljung, G. M. (1986). Diagnostic Testing of Univariate Time Series Models. *Biometrika* 73, 725–30.
- Ljung, G. M. and G. E. P. Box (1978). On a Measure of Lack of Fit in Time Series Models. *Biometrika* 65, 297–303.
- Maravall, A. (1983). An Application of Nonlinear Time Series Forecasting. *Journal of Business & Economics Statistics* 1, 66–74.
- McLeod, A. I. (1978). On the Distribution of Residuals Autocorrelations in Box-Jenkins Models. *Journal of the Royal Statistical Society, Ser. B* 40, 296–302.
- McLeod, A. I. and W. K. Li (1983). Diagnostic Checking ARMA Time Series Models Using Squared-residuals Autocorrelations. *Journal of Time Series Analysis* 4, 269–273.
- Monti, A. C. (1994). A Proposal for Residual Autocorrelation Test in Linear Models. *Biometrika* 81, 776–80.
- Priestley, M. B. (1989). *Spectral Analysis and Time Series, Vol.1*. Academic Press Inc., San Diego.
- Tiao, G. C. and R. S. Tsay (1994). Some Advances in Non-linear and Adaptive Modeling in Time Series. *Journal of Forecasting* 13, 109–131.
- Tsay, R. S. (1986). Nonlinearity Test for Time Series. *Biometrika* 73, 461–466.
- Velilla, S. (1994). A Goodness-of-fit Test for Autoregressive Moving-average Models Based on the Standardized Sample Spectral Distribution of the Residuals. *Journal of Time Series Analysis* 15, 637–47.

Table 1: Empirical means of D_m , D_m^* , Q_{LB} and Q_{MT} with 10,000 replications of a white noise process. Sample sizes are $n=50, 100$ and 500 and $m=7, 15$ and 20 . In the last row the theoretical mean of each statistic is shown.

Empirical Mean												
n	m=7				m=15				m=20			
	D_m	D_m^*	Q_{LB}	Q_{MT}	D_m	D_m^*	Q_{LB}	Q_{MT}	D_m	D_m^*	Q_{LB}	Q_{MT}
50	3.68	4.10	7.16	7.29	6.67	7.97	15.29	15.31	8.21	10.32	20.36	20.15
100	3.85	4.06	7.06	7.14	7.32	7.95	15.17	15.24	9.32	10.32	20.21	20.25
500	4.00	4.05	7.04	7.05	7.88	8.00	15.03	15.07	10.27	10.46	20.00	20.04
Theoretical Mean												
∞	4	4	7	7	8	8	15	15	10.5	10.5	20	20

Table 2: Empirical variance of D_m , D_m^* , Q_{LB} and Q_{MT} with 10000 replications of a white noise process. Sample sizes are $n=50, 100$ and 500 and $m=7, 15$ and 20 . In the last row the theoretical variance of each statistic is shown.

Empirical Variance												
n	m=7				m=15				m=20			
	D_m	D_m^*	Q_{LB}	Q_{MT}	D_m	D_m^*	Q_{LB}	Q_{MT}	D_m	D_m^*	Q_{LB}	Q_{MT}
50	4.42	5.38	16.57	14.24	6.80	9.49	44.86	28.81	7.71	11.86	65.54	36.48
100	5.04	5.53	15.03	13.92	8.62	9.92	37.95	30.04	10.44	12.42	53.67	39.53
500	5.82	5.92	14.50	14.12	10.50	10.77	30.82	29.37	13.38	13.79	42.10	39.58
Theoretical Variance												
∞	5.71	5.71	14	14	11.02	11.02	30	30	14.35	14.35	40	40

Table 3: Significance levels of Q_D , Q_{LB} and Q_{MT} under an AR(1) model.

ϕ_1	m=10			m=15			m=20		
	D_m^*	Q_{LB}	Q_{MT}	D_m^*	Q_{LB}	Q_{MT}	D_m^*	Q_{LB}	Q_{MT}
$\alpha = 0.05$									
0.10	0.055	0.053	0.056	0.054	0.057	0.054	0.055	0.062	0.054
0.30	0.053	0.052	0.053	0.052	0.057	0.054	0.053	0.062	0.049
0.50	0.052	0.055	0.052	0.049	0.057	0.046	0.047	0.058	0.046
0.70	0.054	0.057	0.053	0.050	0.060	0.048	0.050	0.069	0.048
0.90	0.050	0.054	0.043	0.042	0.059	0.039	0.041	0.061	0.039
$\alpha = 0.01$									
0.10	0.009	0.013	0.011	0.009	0.016	0.011	0.010	0.019	0.011
0.30	0.010	0.012	0.011	0.009	0.017	0.011	0.009	0.021	0.010
0.50	0.008	0.012	0.008	0.007	0.016	0.007	0.007	0.019	0.008
0.70	0.010	0.015	0.010	0.008	0.015	0.010	0.009	0.021	0.007
0.90	0.011	0.013	0.009	0.009	0.017	0.009	0.009	0.021	0.009

Table 4: Power level of Q_D , Q_{LB} and Q_{MT} when data are fitted by AR(1) and MA(1) models under various alternative ARMA(2,2) models

(a) Fitted by AR(1) model										
M	ϕ_1	ϕ_2	θ_1	θ_2	m=10			m=20		
					D_m^*	Q_{LB}	Q_{MT}	D_m^*	Q_{LB}	Q_{MT}
1	-	-	-0.50	-	0.415	0.234	0.278	0.299	0.189	0.178
2	-	-	-0.80	-	0.987	0.751	0.959	0.972	0.590	0.855
3	-	-	-0.60	0.30	0.994	0.762	0.983	0.987	0.655	0.941
4	0.10	0.30	-	-	0.597	0.421	0.412	0.452	0.364	0.296
5	1.30	-0.35	-	-	0.807	0.620	0.605	0.649	0.540	0.415
6	0.70	-	-0.40	-	0.781	0.542	0.609	0.637	0.428	0.415
7	0.70	-	-0.90	-	1.000	0.982	0.998	0.998	0.905	0.992
8	0.40	-	-0.60	0.30	0.999	0.836	0.998	0.997	0.697	0.965
9	0.70	-	0.70	-0.15	0.216	0.175	0.169	0.182	0.173	0.111
10	0.70	0.20	0.50	-	0.858	0.759	0.763	0.781	0.658	0.621
11	0.70	0.20	-0.50	-	0.599	0.324	0.384	0.447	0.288	0.260
12	0.90	-0.40	1.20	-0.30	0.988	0.694	0.971	0.970	0.555	0.893
(b) Fitted by MA(1) model										
	ϕ_1	ϕ_2	θ_1	θ_2	m=10			m=20		
					D_m^*	Q_{LB}	Q_{MT}	D_m^*	Q_{LB}	Q_{MT}
13	0.50	-	-	-	0.366	0.295	0.265	0.287	0.243	0.189
14	0.80	-	-	-	0.993	0.984	0.980	0.987	0.969	0.939
15	1.10	-0.35	-	-	1.000	1.000	0.999	0.999	0.986	0.988
16	-	-	0.80	-0.50	0.988	0.851	0.940	0.953	0.727	0.838
17	-	-	-0.60	0.30	0.674	0.400	0.476	0.540	0.337	0.340
18	0.50	-	-0.70	-	0.957	0.888	0.876	0.888	0.801	0.736
19	-0.5	-	0.70	-	0.957	0.893	0.876	0.908	0.807	0.758
20	0.30	-	0.80	-0.50	0.859	0.583	0.743	0.765	0.468	0.556
21	0.80	-	-0.50	0.30	0.992	0.980	0.960	0.973	0.968	0.916
22	1.20	-0.50	0.90	-	0.706	0.467	0.696	0.708	0.393	0.570
23	0.30	-0.20	-0.70	-	0.426	0.278	0.280	0.306	0.233	0.208
24	0.90	-0.40	1.20	-0.30	0.965	0.780	0.923	0.932	0.638	0.822

Table 5: Empirical frequency of rejection of the null hypothesis of linearity, when the fitted model is an AR(p), and n=204 with 1,000 replications. Nominal significance level, $\alpha = 0.05$.

	m=7			m=12			m=24		
	D_m^*	Q_{LB}	Q_{MT}	D_m^*	Q_{LB}	Q_{MT}	D_m^*	Q_{LB}	Q_{MT}
Model 1	0.157	0.114	0.120	0.126	0.100	0.091	0.091	0.094	0.089
Model 2	0.566	0.497	0.480	0.511	0.426	0.394	0.406	0.289	0.277
Model 3	0.960	0.900	0.914	0.940	0.854	0.826	0.870	0.740	0.70
Model 4	0.914	0.860	0.830	0.874	0.805	0.760	0.788	0.686	0.614

Table 6: Number of times the statistics D_m^* , Q_{LB} and Q_{MT} find structure in the residuals $\hat{\varepsilon}_t$, (with $\alpha = 0.05$) considering 12, 24 and 36 lags.

	m=12			m=24			m=36		
n	D_m^*	Q_{LB}	Q_{MT}	D_m^*	Q_{LB}	Q_{MT}	D_m^*	Q_{LB}	Q_{MT}
100	0.045	0.036	0.046	0.043	0.044	0.045	0.025	0.056	0.032
200	0.054	0.051	0.053	0.045	0.050	0.056	0.037	0.059	0.044
500	0.054	0.048	0.050	0.046	0.054	0.046	0.040	0.059	0.051

Table 7: Power of the statistics when an AR(2) model is fitted and the true one is generated according to TAR(2) model.

	m=12			m=24			m=36		
n	D_m^*	Q_{LB}	Q_{MT}	D_m^*	Q_{LB}	Q_{MT}	D_m^*	Q_{LB}	Q_{MT}
100	0.162	0.115	0.109	0.089	0.076	0.077	0.075	0.075	0.055
200	0.296	0.228	0.213	0.216	0.17	0.178	0.18	0.146	0.122
500	0.648	0.527	0.524	0.556	0.442	0.415	0.499	0.369	0.342

Table 8: Results of the nonlinearity test for the Sunspot data, when the AR(9) model is fitted.

m	$D_m^*(\hat{\varepsilon}_t^2)$	$D_m^*(\hat{\varepsilon}_t^2)/d_C$	$Q_{LB}(\hat{\varepsilon}_t^2)$	$Q_{LB}(\hat{\varepsilon}_t^2)/d_C$	$Q_{MT}(\hat{\varepsilon}_t^2)$	$Q_{MT}(\hat{\varepsilon}_t^2)/d_C$
7	17.35	1.93	21.13	1.50	19.15	1.36
12	18.52	1.53	22.9	1.08	21.17	1.01
24	21.01	1.05	27.07	0.74	26.56	0.73

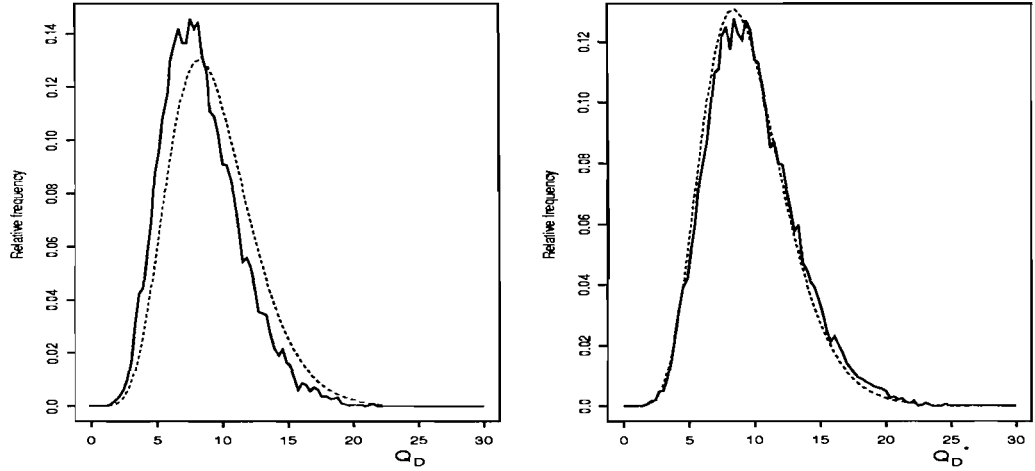


Figure 1: Monte Carlo distributions of D_m and D_m^* (calculate with the standardized autocorrelation), dashed line. 10000 replications of the AR(1) model with $\phi_1 = 0.5$, when $n=100$ and $m=20$; solid line is the approximation of equation (12) by $a\chi_b^2$ ($a = 0.545$, $b = 17.44$).